

Risks of Hyperthermia Associated with Hot Tub or Spa Use by Pregnant Women

Technical Report prepared by:
Christina Chambers, Ph.D., M.P.H.
Under Contract to
U.S. Consumer Product Safety Commission

Contract CPSCS041343

The recommendations in this report are those of the contractor and may not reflect the views of the CPSC staff. This report is being posted on the CPSC web site for public comment. CPSC staff will develop its recommendations after reviewing the public comments. The staff's draft recommendations will be posted on the CPSC web site under Voluntary Standards/Research Reports before being finalized.

This report has not been reviewed or approved by the Commission and may not reflect the views of the Commission.

Table of Contents

I.	Introduction and Purpose of the Report.....	3
II.	Exposures and Outcomes – Primary and Secondary	3
III.	Principles of Teratology Used for Review	4
IV.	Approach to the Literature Review	4
V.	Summary of Evidence in Key Areas of Research	6
	a. Animal data	6
	b. Human data.....	7
	i. Maternal fever or febrile illness	7
	ii. Maternal hot tub or sauna use	10
	c. Generalizability of published data to hot tub or spa use in pregnancy	12
	d. Potential risk modifiers	16
VI.	Relevance of Data to Technical Specifications for Hot Tub or Spa Construction	17
VII.	Conclusions and Recommendations	17
VIII.	References	18

I. Introduction and Purpose of the Report

Birth defects are the leading cause of infant mortality in the United States, accounting for more than 20% of all infant deaths. Of about 120,000 U.S. babies born each year with a birth defect, 5,600 die during their first year of life. In addition, birth defects are the fifth-leading cause of years of potential life lost and contribute substantially to childhood morbidity and long-term disability. Although the cause of about 70% of all birth defects is unknown, some defects are known to be caused by maternal exposures to certain environmental agents during pregnancy (Centers for Disease Control and Prevention, 2005).

Environmental causes of birth defects and other adverse reproductive outcomes (also known as teratogens) are by definition preventable if the environmental cause that confers increased risk for these negative outcomes can be avoided. Maternal hyperthermia or elevated maternal body temperature has been identified in several studies as an environmental cause of major birth defects, as well as a wider pattern of adverse reproductive outcomes. The most commonly reported category of major congenital anomalies linked with elevated maternal body temperature is central nervous system defects that involve improper closure of the neural tube. These defects may be incompatible with fetal or newborn life. If the fetus survives, neural tube defects almost always require surgical correction, and the associated morbidity can represent lifelong disability.

This report was requested by the U.S. Consumer Product Safety Commission staff to provide an overview and critical assessment of the literature with respect to the technical specifications and typical use of hot tubs or spas by women of reproductive age. The specific focus of this review is to assess the evidence for an increased risk of neural tube defects in the offspring of women who use hot tubs or spas while they are pregnant.

II. Exposures and Outcomes – Primary and Secondary

The primary exposure that is the focus of this critical review is hot tub or spa use. However, since there is a paucity of data on this specific exposure in human pregnancy, studies that focus on maternal hyperthermia resulting from fever associated with illness, and to a lesser extent sauna use will be cited and used for extrapolation. Animal studies that involve manipulation of environmental temperatures will also be cited. However, human studies that focus on other potential sources of elevated maternal body temperature, such as electric blanket use, workplace or occupational exposures, exercise-induced hyperthermia, and localized hyperthermia associated with diagnostic ultrasound were not considered as primary to this review – and may be mentioned only secondarily.

The endpoint that is the focus of this critical review is the group of congenital anomalies that comprise neural tube closure defects (NTDs). The number of studies, and to some

extent the most compelling human evidence, exists for this outcome of pregnancy in association with maternal hyperthermia. However, secondary endpoints that will be reviewed to a lesser extent include spontaneous abortion or stillbirth, and other major congenital anomalies such as heart defects.

For purposes of this review, temperatures will be expressed in degrees Celsius (C); however, a Fahrenheit conversion table is provided for reference in the appendices.

III. Principles of Teratology Used for the Review

Established principles of teratology will be utilized as a basis for discussion of findings of published studies cited in this review. These principles as set forth by Wilson and Fraser (1977) include:

1. Susceptibility to teratogenesis depends on the genotype of the conceptus and the manner in which this interacts with environmental factors.
2. Susceptibility to teratogenic agents varies with the developmental stage at the time of exposure.
3. Teratogenic agents act in specific ways (mechanisms) on developing cells and tissues to initiate abnormal embryogenesis (pathogenesis).
4. The final manifestations of abnormal development are death, malformation, growth retardation, and functional disorder.
5. The access of adverse environmental influences to developing tissues depends on the nature of the influences (agent).
6. Manifestations of abnormal development increase in degree as dosage of the teratogenic agent increases from the no-effect to the totally lethal level.

With respect to the primary endpoint for this review, it is relevant to the second principle that the timing of neural tube closure in the human embryo is between the second and fourth week of development following conception. For other major defects, such as heart defects or oral clefts, the critical window for exposure may extend from two weeks post-conception to nine weeks, or near the end of the first trimester.

IV. Approach to the Literature Review

We conducted 12 searches of Medline/PubMed using the National Library of Medicine's MESH terms. All searches were set for Publication Date from 1960 to 2004.

Selected articles were exported in Medline format and imported to an EndNote library with duplicates excluded. The following searches were performed:

1. Abnormalities and Heat/adverse effects and Pregnancy Trimester, First

2. Abnormalities and Fever/complications and Pregnancy Trimester, First
3. Abnormalities and Hyperthermia, Induced/adverse effects and Pregnancy Trimester, First
4. Embryonic and Fetal Development and Heat/adverse effects
5. Embryonic and Fetal Development and Fever/complications
6. Embryonic and Fetal Development and Hyperthermia, Induced/adverse effects
7. Abnormalities and Heat/adverse effects and Pregnancy
8. Abnormalities and Fever/complications and Pregnancy
9. Abnormalities and Hyperthermia, Induced/adverse effects and Pregnancy
10. Abnormalities and Heat/adverse effects and Pregnancy Complications
11. Abnormalities and Fever/complications and Pregnancy Complications
12. Abnormalities and Hyperthermia, Induced/adverse effects and Pregnancy Complications

Additional searches were conducted using key publications to determine if other publications that had included the key paper in citations (Web of Science) were relevant. Research librarian assistance was secured to review alternative databases to determine if relevant publications may have appeared in the literature but were not indexed in Medline/PubMed. References in the English language literature, or with English abstracts, were downloaded into an Endnote library, and all abstracts were reviewed in detail for relevance to the topic. The full text articles for those abstracts selected for retention in the library were then obtained and reviewed in their entirety.

The searches, initially conducted on November 2, 2004, yielded 361 articles (including duplicates) of which 155 were determined to be of interest and are included in the Endnote library. Within the library, 104 articles are labeled "Full Text Review" and were critically reviewed in their entirety; 14 are labeled "Foreign Language" for which only abstracts were reviewed; 36 are labeled "Additional Articles of Interest" for which only abstracts were reviewed, and 1 article is listed as "Unavailable" and could not be obtained in full text form. The subset of articles in this library which have been cited in this technical report are referenced at the end of the report, and hard copies of the full text of these manuscripts are supplied with the report.

V. Summary of Evidence in Key Areas of Research

a. Animal Data

Dating back to the 1960's, Marshall Edwards, an Australian veterinarian, first noted that pregnant guinea pigs that were exposed coincidentally to unusually high environmental temperatures had a remarkably high incidence of adverse pregnancy outcomes. This astute observation led to numerous experimental animal studies designed to confirm the association, and to further define the relationship between the dose or height of maternal temperature elevation and the length of exposure required to induce abnormalities (Edwards et al, 1995; Graham et al, 1998).

As summarized in Table 1 attached, a wide variety of species including guinea pig, rat, rabbit, mouse, hamster, bonnet monkey, and sheep have demonstrated susceptibility to elevated maternal body temperature in the induction of a wide variety of adverse reproductive outcomes. Increased risks for these outcomes noted in animal studies are consistent with each of Wilson's principles of teratology. Of particular importance, susceptibility to specific outcomes with the experimental manipulation of temperature is consistent with the second principle – that susceptibility depends on the developmental stage at the timing of exposure. Thus, for example, the highest risk for neural tube defects in the guinea pig is with exposure at day 13 (Cawdell-Smith et al, 1992).

Also of critical importance is the consistent finding in the animal literature that, above a certain threshold dose, the combination of increasing dose and/or increasing duration of exposure to maternal hyperthermia is related to increased risk for abnormal embryonic development. This is consistent with the sixth principle of teratology – that manifestations of abnormal development increase in degree as dosage of the teratogenic agent increases from the no-effect level to the totally lethal level. Thus, for example, an increased risk for neural tube defects, facial clefting, and other defects is seen in the Sprague-Dawley rat only at core temperature increases above 2.0 degrees C over baseline. Furthermore, an increase in core temperature of 3.0 degrees C above baseline is sufficient for teratogenesis for exposures lasting at least 20 minutes, whereas an increase in core temperature of 4.0 degrees C above baseline is sufficient for teratogenesis after only 5 minutes duration of exposure (Germain et al, 1985).

In addition to the apparent threshold of about 2 degrees C elevation in body temperature across animal species, it is important to note that the increase in core body temperature that is induced by an environmental exposure to a heat source in animal studies is maintained for a period of time beyond removal of the source of heat. Thus repeated exposures with short intervals away from the heat source may not allow the mother to return to normal body temperature and may exacerbate the effects of hyperthermia.

The first principle of teratology indicates that susceptibility to teratogenesis depends on the genotype of the conceptus and the manner in which this interacts with environmental factors. This principle often is used to explain findings in animal studies that are inconsistent across species and inconsistent with human data regarding the same exposure. However, across species there is remarkable consistency in the animal literature in the spectrum of congenital anomalies that are seen with increased frequency, including neural tube defects, and the magnitude of temperature elevation required to induce a measurable excess of these defects. These findings suggest cross-species similarities with respect to susceptibility, but nevertheless, as the first principle of teratology indicates, human data is required to demonstrate human susceptibility to teratogenesis, and especially to define the dose and duration required to confer increased risk for congenital anomalies.

b. Human Data

i. Maternal fever or febrile illness

The earliest human evidence in support of Edwards' initial animal data was based on a case series published by Smith et al in 1978. The case series identified a pattern of defects associated with exposure between four and six weeks' gestation and another pattern of defects associated with exposure between 7 and 16 weeks' gestation. These patterns included microphthalmia, oral clefts, limb abnormalities, and central nervous system problems. The sources of hyperthermia exposure were primarily high fever – on the order of 38.9 ° C for at least 1-2 days in duration, although there was one case included in the series of hyperthermia exposure in association with sauna use. This series lent compelling support to human susceptibility to the teratogenic effects of maternal hyperthermia because a specific pattern of defects was noted, with specific timing of exposure and an apparent threshold dose – findings consistent with the 2nd, 3rd, 4th and 6th principles of teratology. Of particular significance was that 12 of the 13 cases in the series were associated with maternal febrile illness, while only one case was associated with maternal sauna use.

Subsequently, several case control studies appeared in the literature testing the hypothesis that maternal fever is associated with birth defects. Case control studies in birth defects epidemiology are typically constructed with cases that manifest a single major malformation, or category of malformations. Cases are then compared to controls consisting typically of non-malformed infants. The limitations of case control studies in establishing cause and effect are well described and include the possibility of recall bias (that mothers of malformed infants may be motivated to recall differently than mothers of normal infants, and that both case and control mothers may have poor or inaccurate recall about the timing, duration and magnitude of any

temperature elevation that occurred months to years prior to the study interview). These limitations are further amplified with an exposure like hyperthermia which is often difficult or impossible to validate from any source other than maternal report (Christensen, 2002). In observational studies which rely on maternal report of a febrile illness that has already occurred, pregnant women may not have taken and recorded their own temperature, may not have taken and recorded it frequently over the course of an illness, and if they have measured temperature elevations, they may never have reported these values to a health care provider, thereby obstructing possible validation through the medical record. Similarly, with other sources of hyperthermia incurred in a recreational setting, such as hot tub or spa use, in observational studies, it is unlikely that women recalling such exposures would have recorded their body temperature and the duration of the exposure in any manner suitable for validating maternal recall.

As summarized in Table 2 attached, the human data taken as a whole are not comprehensive, consist of observational studies which may or may not be able to appropriately control for confounding, and are limited in ability to tease apart the potential effect of hyperthermia from the potential effect of an underlying disease leading to fever or the medications used to treat it. Furthermore, a major limitation of these retrospective case-control studies is the typical inability to identify the specific timing of the exposure and the specific magnitude and duration of the elevated body temperature.

Nevertheless, several case-control studies demonstrate an association with NTDs and maternal fever, and some well-constructed case-control studies have shown this association with statistical significance (Miller et al, 1978; Layde et al, 1980; Sandford et al, 1992; Lynberg et al, 1994). The fact that these studies are consistent with the animal literature, and that other case-control studies have demonstrated an association with first trimester maternal fever and increased risk for other defects including heart defects, and Hirschprung Disease (Tikkanen et al, 1991; Botto et al, 2001 and 2002; Lipson et al, 1988), all lead to the conclusion that the association in humans between maternal hyperthermia and birth defects is causal.

A very small number of prospective cohort studies have addressed these same hypotheses. Prospective studies generally offer the advantage of more recent and precise maternal recall about exposures, and can avoid to some extent the potential bias of retrospective studies by collecting information about exposure before the known outcome of pregnancy. However, the contribution these kinds of study designs can make to evaluating the effect of hyperthermia on risk for NTDs is usually limited. This is primarily a sample size issue. In a population or clinic based sample, the number of pregnant women who have had a documented elevated body temperature of at least 2 degrees C above baseline that has endured for a documented 24 to 48 hour period, and that has occurred during the very short and early critical period for

neural tube closure is likely to be quite small. Furthermore, even though neural tube defects are one of the most common types of major congenital malformations, a baseline risk of 1 per 1,000 live births translates into a necessary sample size of exposed and unexposed mothers in the thousands in order to have sufficient power to rule out a doubling of risk for that specific defect. Such sample sizes are usually well beyond the bounds of attainability and cost in most cohort study designs.

Nevertheless, four cohort studies bear mentioning. Clarren et al (1979) used the Collaborative Perinatal Project cohort of over 50,000 mother child pairs to identify 178 women who reported a fever of at least 38.9 degrees C or higher on at least one occasion in the first trimester in association with an illness (38.9 degrees C is approximately 1.9 degrees C above the average baseline temperature in humans). No significant associations were found for all major malformations combined or any other adverse outcome. Kleinfelder et al (1980) looked at any malformation or other developmental problem up to three years of age in a cohort of 7,870 pregnancies. An association between flu in pregnancy and abnormal muscle tone in the offspring was the only significant finding. Little et al (1991) reported on 54 women exposed to a fever of 38.3 degrees C or higher for 24 hours or more in the first trimester, and found significant increased risks for diastasis recti or herniated umbilicus compared to unexposed matched controls.

Finally, Chambers et al (1998) identified 115 women exposed to a fever of 38.9 degrees C or higher for at least 24 hours in the first trimester and found no significant differences in overall proportion of malformed infants compared to controls who reported a lower fever or a fever of shorter duration, and also compared to controls who did not report any fever. However, the authors did note two NTDs among infants of the 34 women who reported a high fever during the critical period for neural tube closure. This finding was not statistically significant but far exceeded the number of expected neural tube defects in a sample of this size. The specific finding of a greater number of NTDs than expected provided not only evidence consistent with previous case-control studies, but also provided some measure of the magnitude of the risk for an NTD with prenatal exposure to a fever of the specified elevation and duration during the critical window of susceptibility. Furthermore, as part of this cohort study design, a subset of live born infants received a specialized dysmorphological examination to evaluate children for minor and major structural defects. In this study an excess of specific minor malformations was noted among infants exposed to high fever, and importantly, these minor malformations were consistent with the spectrum of defects and timing of exposure identified by Smith and others in the previously published case series. Thus, this study was the first prospective evaluation of maternal fever that was able to confirm findings from previous case-control studies and case series with respect to the magnitude and duration of elevated maternal body temperature in relation to an increased risk for NTDs as well as a specific

pattern of malformation. The consistency of findings across these different study designs further lends support to the causal nature of the relationship between maternal fever and specific birth defects.

Other human studies have focused on a range of adverse outcomes. Kline et al (1985) used a case control methodology to show a significant association between late pregnancy spontaneous abortion (<28 weeks gestation) and a maternal fever of 37.8 degrees C or higher. Furthermore, despite the case-control design, Kline and colleagues were able to effectively rule out maternal recall bias by comparing maternal recall of fever in mothers of abortuses with normal karyotype to those with abnormal karyotype, prior to mothers being made aware of the karyotype status of their pregnancy. In the latter situation, there was no excess of maternal fever reported. This is consistent with the presumed cause of the miscarriage as it is expected that a high proportion of fetuses with abnormal chromosomes will spontaneously abort. However, there was an excess of reported maternal fever in mothers whose abortuses had normal karyotype. In contrast, Andersen et al (2002) using the Danish national cohort study consisting of over 24,000 pregnancies, did not find any association between maternal fever in the first 16 weeks of pregnancy and spontaneous abortion or stillbirth. Of note, in this study, more than 18% of the entire sample of women reported a fever occurring sometime in the first 16 weeks of pregnancy, although many did not recall how high the fever was or how long it lasted.

In summary, the consistency of findings across case-control studies taken as a whole supports the conclusion that maternal febrile illness increases the risk for major congenital defects, with most evidence in support of the risk for NTDs. Although cohort studies have produced conflicting results with respect to major defects, typically they have been underpowered to test the hypothesis that maternal fever causes NTDs. The one cohort study that did produce results consistent with both case-control and case series findings had the advantage of a specialized evaluation of a subset of live born infants – therefore improving the sensitivity of the study to detect differences between exposed and unexposed infants, despite relatively small sample size. These data taken as a group support the conclusion that maternal hyperthermia contributes to an increased risk for NTDs as well as other structural defects that represent a characteristic pattern.

ii. Maternal hot tub or sauna use

Few studies have contributed information to knowledge about the risks associated with hot tub or sauna use. Although animal data would suggest that a sufficient increase in core body temperature, regardless of the source, confers risk, evidence supporting that exposure to hot tubs or saunas is teratogenic is sparse.

Several relatively small case control-studies have incorporated “other environmental sources of heat” in study questionnaires. Halperin and Wilroy (1978) reported on a small case control study (48 NTD cases and matched controls) and found 3 case mothers who reported an elevated body temperature vs. 1 control. One of the case mothers reported sauna use on four occasions in the fifth week of gestation with a report of one occasion where the temperature was 43 degrees C. Chance et al (1978) in a similarly small case-control study (43 cases of NTD and 63 controls) reported that no case mothers recalled using a very hot sauna for 15 minutes or more during the period of neural tube closure. Lipson et al (1988) found no association between non-fever sources of hyperthermia and Hirschsprung Disease. Similarly, Tikkanen et al (1991) in a case control study in Finland found maternal fever during early pregnancy to be more common among 573 case mothers of infants with heart defects, but found no association with sauna bathing, workplace temperatures, or temperature of the environment. In another Finnish study, Saxen et al (1982) compared sauna bathing habits in the mothers of 100 consecutive cases of infants with defects of the central nervous system and 202 control mothers whose infants had orofacial clefts and found no differences between the groups.

On the other hand, Miller et al (1978) reported on 63 NTD cases and 64 controls, and found that hyperthermia was statistically significantly associated with being a case mother. Of the six instances of hyperthermia, two mothers possibly had experienced hyperthermia related to sauna use near the time of neural tube closure. Fisher and Smith (1981) reported on 17 infants with encephalocele compared to matched children with Down Syndrome and found that maternal fever of at least 1.9 degrees C above baseline was reported at the critical period in four cases and no controls. Two further cases in this series included one child whose source of hyperthermia was a one-hour Japanese bath (not otherwise described). Sandford et al (1992) reported on 44 NTD cases and 44 matched controls and found a statistically significant association with maternal report of “hot baths” in the first gestational month and NTDs.

Li et al (2003a) examined spontaneous abortions in a Kaiser cohort study in which women were interviewed during pregnancy about hot tub use. Although more than half of mothers were interviewed after the spontaneous abortion had already taken place, the authors reported a 2 fold increased risk for spontaneous abortion with hot tub or whirlpool exposure after conception. The authors further noted an increase in risk with increased frequency of use and increasing temperature of the water.

Finally, the single most concerning study regarding hot tub use was published by Milunsky et al in 1992. Based on an ongoing prospective cohort study in which women were recruited primarily through private obstetric offices in New England, 23,491 pregnant women were enrolled almost exclusively between

15 and 20 weeks' gestation at the time of prenatal screening or diagnosis with maternal serum alpha fetoprotein or amniocentesis. Maternal telephone interviews regarding pregnancy exposures were conducted by trained nurses. Among these women, 737 were excluded due to missing heat exposure information (n=1) or missing outcome information (n=736). Among the remaining subjects, a total of 5,566 women were exposed to at least one heat source including either hot tub, sauna or electric blanket in the first two months of pregnancy or fever greater than or equal to about 38 degrees C in the first three months of pregnancy. There were 1,254 women in the sample who reported using a hot tub or whirlpool in the first two months of pregnancy (for 55 women hot tub exposure was unknown). Forty-nine pregnancies in the entire cohort were identified as involving a fetus with a neural tube defect. In crude analysis, an almost 3 fold, statistically significantly elevated increased risk was found for hot tub or whirlpool use in the first two months of pregnancy and NTDs compared to women with no heat exposure. This relationship held up even after adjustment for maternal age, family history of NTDs, use of folic acid in the first six weeks of pregnancy, and exposure to other heat sources (adjusted relative risk 2.8; 95% confidence interval 1.2-6.5). This translates into a risk of approximately 2-3 per 1,000 NTDs in women who have used hot tubs in the first two months of pregnancy, or an excess of 1-2 NTDs per 1,000 exposed women over the baseline rate of approximately 1 per 1,000 births. In contrast, no statistically significant independent association was found for sauna use, fever, or electric blanket use. Furthermore, the authors found a six fold, statistically significant increased risk for NTDs among women who had at least two heat exposures to any of the following during the first two to three months of pregnancy: hot tub, sauna or fever, compared to those who reported no heat exposure at all.

c. Generalizability of Published Data to Hot Tub or Spa Use in Pregnancy

As indicated above, the only published study that had a reasonable sample size of pregnant women exposed to hot tub use during the approximate critical period for neural tube closure is the Milunsky study. To reiterate, the association in this paper was statistically significant with an adjusted relative risk of 2.8, 95% confidence interval 1.2-6.5. With this study taken on its own merit, there is evidence to suggest based on the lower bound of the confidence interval that there is at least a 20% increased risk for NTDs if a pregnant woman uses a hot tub at least once in the first two months of gestation. One limitation of this study, although described as prospective, is that more than half of all women were interviewed about exposure to hyperthermia after the results of prenatal diagnosis were known, introducing the potential for recall bias. Another important limitation of this study is that details about hot tub use were not collected. Thus, the frequency of use was documented, but not the length of time in the tub, the level of immersion in the water, or the water temperature at the time the hot tub was used. Because a statistically significant association was found, despite the fact that data on dosage and duration of hyperthermia

exposure were missing and the timing of exposure encompassed a broader time period than the short critical window for development of NTDs, it is possible that the association would be stronger (i.e., the relative risk higher) if that information were known. The effect of random misclassification (i.e., some women being classified as “exposed” when their exposure was not during the critical period, or was too brief to confer risk, etc., is thought to erroneously dull the estimate of the relative risk due to “noise” in the data. However, it is not possible to be sure in an observational study such as this. Therefore, even though these specific findings may be valid, it is difficult to apply them to the individual situation, i.e., to suggest that on the average every woman who uses a hot tub for any length of time and at any temperature during the first two months of pregnancy is at increased risk of having a child with an NTD.

Furthermore, with any finding in an observational study such as this, it is difficult to be sure that this association represents a cause and effect relationship. This is true because it is not ethically reasonable to conduct the definitive study which would be a randomized clinical trial that, by design, controls for measured and unmeasured confounding. Furthermore, there is no series of well-designed human studies with varying research designs that have come to the same conclusion as the Milunsky study.

As mentioned above, the only other study with a significantly positive finding for NTDs in association with a perhaps comparable exposure was Sandford et al’s 1992 case control study which found a higher frequency of hot baths in the first gestational month reported by mothers of infants with NTDs compared to matched controls. Again, the temperature and duration of exposure was not available in this study.

Similarly, the recent cohort study published by Li et al (2003a), showing a doubling of risk for spontaneous abortion with hot tub or whirlpool exposure, included no information on temperature of the water or duration of exposure. This study could also have involved recall bias in that more than half of women who had spontaneous abortions were interviewed about exposures after the event had already occurred.

To summarize, the limited data available that directly bear on the exposure of interest – hot tubs or spas – suggest an association with NTDs. To date, the best estimate of the risk associated with hot tub use during the critical window of NTD closure is a minimum of a 20% increased risk, with a point estimate of a 280% increased risk. This estimate translates to an increase over the baseline risk of 1 per 1000 to a risk of 2-3 per 1000 for an exposed woman to have an NTD-affected pregnancy.

With this limited amount of evidence directly related to hot tub or spa use, it is relevant to view this evidence in the context of data on other sources of hyperthermia.

Some have suggested that the association between fever-related hyperthermia and NTDs could be attributable to the maternal disease and/or the medications used to treat the illness. Yet, the association has been noted in studies where the maternal illness has varied from flu to kidney infection, and there is relatively consistent data supporting a dose-response relationship with higher fever, especially fevers resulting in temperature elevations of approximately 2 degrees C above baseline. The consistency of these findings both in terms of the threshold dose, and the effect of temperature elevation in the absence of infection as noted in numerous animal studies, supports the concept that maternal hyperthermia itself is at least in part causally related to the risk for NTDs.

With respect to non-fever related sources of hyperthermia, it has been argued that maternal sauna bathing has been associated with NTDs only in isolated instances (Smith et al, 1978; Halperin et al, 1978), and that in countries where sauna bathing is frequent among pregnant women, the rates of NTDs are not notably higher. For example, Tikkanen et al (1991) reported an association with maternal fever and heart defects in the offspring of Finnish women, but could document no such association with sauna use. In fact, approximately 75% of all cases and controls in this study sauna bathed one or two times per week, and there were no differences between cases and controls on the frequency of use or timing of exposure. In a survey study in the same population, Uhari et al (1979) questioned women about their sauna use during pregnancy and found that 84% of women reported sauna bathing in the first half of pregnancy. It may be that these sauna users are particularly tolerant of the heat exposure, limit their exposure to brief periods, or reduce their core body temperature quickly by taking measures to cool down.

With respect to self-modulation of maternal exposure to hyperthermia, it has been suggested that sauna bathers in Finland may reduce body temperature between sauna sessions by immersion in snow or cold water (Lipson et al, 1993). Unlike fever-related hyperthermia that may not be under the control of the pregnant woman, hot tub or sauna bathing may pose less of a risk because the pregnant woman can self-limit her exposure, especially if the temperature becomes uncomfortable or unbearable. Furthermore, women who use hot tubs can control the level of immersion and the temperature of the water, although it is important to recognize that one of the normal mechanisms for body temperature regulation, i.e., perspiration, is limited in the hot tub setting (Edwards et al, 1995).

Relevant to the possible differences in tolerance to heat exposures with hot tub or sauna use, a study conducted by Harvey et al (1981) examined the effect of various hot tub water temperatures and sauna temperatures on increasing core body temperature and ability to tolerate these temperature elevations in non-pregnant women. In this study, 20 healthy non-pregnant volunteers immersed in hot tubs heated to 39 to 41 degrees C. At 39 degrees subjects began to leave

the tub because of discomfort after 10 minutes, and only five women could remain in the tub long enough to reach a core body temperature of 38.9 degrees C. These women could remain in the tub for 15 to 25 minutes. At 41.1 degrees C, subjects began to leave in discomfort after 5 minutes. However, six women remained in the tub until their body temperatures reached 38.9 degrees and these women stayed in for 10 to 30 minutes. Following the hot tub experiment, these same 20 subjects were exposed to sauna bathing on a different experimental day. None of the subjects who entered a sauna with an average temperature of 81.4 degrees C could remain in the heated environment long enough to reach a core body temperature of 38.9 degrees C. Of interest, some women in each condition (hot tub or sauna) whose temperatures were measured 10 minutes after the heat exposure showed an increase in core temperature above the final reading in the heated environment. This could suggest that women who leave and re-enter a hot tub with only a short interval in-between (less than 10 minutes) may achieve the same maximum elevated body temperature that would have been achieved by remaining in the heated environment the entire time, i.e., that short intervals between exposures may not be sufficient to reduce body temperature to normal.

Based on this study in non-pregnant women, the authors concluded that it was unlikely that typical recreational exposures to hot tubs or saunas would present a significant risk for NTDs or other hyperthermia-related adverse outcomes because women would self-limit their exposures to a level below the threshold of effect. The authors further suggested that prolonged exposures might pose a risk even when interrupted by short cooling off periods. The guidelines suggested by the authors of this study for maximum temperature and duration of exposure (no more than 15 minutes at a water temperature of 39 degrees C) are aimed at keeping maximum body core temperature below 38.9 degrees C and would be conservative estimates of the limits of safety for pregnant women.

However, there are of course a number of limitations to these conclusions – first, these women were not pregnant – and ability to modulate temperatures effectively may differ during pregnancy. Secondly, some women may be naturally more or less tolerant of heat than the average person, and Harvey et al did note that there was large inter-individual variation in their sample. Therefore, some women may be more or less susceptible or resistant to the effects of environmental sources of hyperthermia. Third, the perception of discomfort may be modulated by other factors commonly associated with a social setting but not involved in an experimental setting. For example, social interactions surrounding hot tub use may distract from or moderate perceptions of discomfort, or may influence the perception of elapsed time; and the use of alcohol before or while immersed in a hot tub may also modulate body temperature, affect perception of discomfort, passage of time, and/or impair decision-making ability. Finally, other pre-existing conditions or events may affect risk. For example, if a woman enters a hot tub after vigorous exercise, or with a mild fever, she may already have an

elevated body temperature that is further aggravated by the hot tub exposure leading her to reach a higher core temperature with a shorter duration of time.

In summary, there is limited specific information linking hot tub or spa use to increased risk for NTDs. However, taken in the context of a relatively large volume of animal data and the association between high maternal fever and NTDs in humans, the potential for hot tub or spa exposure to increase the risk for neural tube defects is likely. It may be that the risks of exposure are limited to the most extreme patterns of use – high temperature and frequent and repeated immersion for lengthy periods of time – analogous to some of the human data on dose and duration of fever.

d. Potential Risk Modifiers

As noted in the first principle of teratology, susceptibility to teratogenesis can be species specific. Within species, genetic susceptibility to an environmental teratogen is suggested due to the fact that not all embryos with a similar exposure are similarly affected by the teratogenic insult. Thus, it is hypothesized that one or more genetic factors may put the embryo at risk of an NTD, and a hyperthermic insult of a certain magnitude and duration is sufficient to push that embryo over the threshold of effect. Genetic susceptibility to environmentally induced neural tube defects has been demonstrated in animal models (Finnell et al, 1986).

Other non-genetic risk factors may be involved; e.g., nutritional deficiencies or excesses may play a role (Ferm and Ferm, 1979; Shin and Shiota, 1999; Li et al, 2003b), and alcohol may act synergistically with hyperthermia to induce NTDs (Shiota et al, 1988). Botto et al, (2002) has shown in a case-control study that the association between maternal febrile illness and one of seven major congenital anomalies is modified by multivitamin use.

In addition, susceptibility to hot tub or spa-induced hyperthermia may be modified by a previous history of having a child with an NTD, maternal age, and other conditions or medications that are associated with increased risk of NTDs, including maternal diabetes, and use of some folate-antagonist drugs such as carbamazepine and valproic acid.

Thus women, who fall into an elevated risk category, or who are already ill with a fever, or have an elevated body temperature from another source, may consider carefully the use of hot tubs or spas in pregnancy, especially during the critical period for neural tube closure.

VI. Relevance of Data to Technical Specifications for Hot Tub or Spa Construction

UL 1563 Electric Spas, Equipment Assemblies, and Associated Equipment, Section 33 (p. 57) indicates that the maximum set point water temperature in the tub is 40 degrees C and that the temperature regulating controls must be adjustable to temperatures below that level. This section also sets tolerance requirements for temperature regulation and control and calibration verification. The temperature regulation control at the maximum setting is required to have a tolerance of no more than plus or minus 3 degrees C. This means that at the maximum level of tolerance in a hot tub or spa that meets specifications, the maximum temperature could reach 43 degrees C.

A backup system consisting of a capillary tube or sensing circuit is required to reduce the risk of loss of temperature control in the event of damage; should this backup system malfunction, there is the potential for temperatures higher than 43 degrees C to be achieved.

In Section 34, (p. 58) it is specified that the temperature limiting controls at the water inlet limit the maximum temperature to 50 degrees C with a tolerance of not more than 3 degrees C. This suggests that under normal conditions, the water at the inlet could reach temperatures as high as 53.0 degrees C and areas surrounding the inlet would be expected to reach relatively higher temperatures than the remaining areas of the hot tub or spa.

These maximum temperature levels, overall and at the inlet, are at or substantially above levels that would be considered at the limits of safety for women in early pregnancy who are immersed in a hot tub or spa for more than a few minutes.

VII. Conclusions and Recommendations

It is widely recognized that in the human embryo the developing central nervous system is exquisitely sensitive to environmental insults. As environmental causes of birth defects are potentially preventable, measures to avoid potentially harmful exposures or minimize risks should be taken. Women of reproductive age are advised by the Centers for Disease Control and Prevention (2005) to take 400 mcg. of supplemental folic acid daily in order to reduce the risk of NTDs should they become pregnant. In addition to following this advice, women who choose to use hot tubs or spas should consider the possibility that they might unknowingly be pregnant. Approximately 50% of pregnancies in the U.S. are unplanned, and the majority of these pregnancies are not recognized until the period of neural tube closure has already passed (Floyd et al, 1999).

Therefore, women who are of reproductive age and have a possibility of being pregnant, based on the findings in non-pregnant women reported by Harvey et al (1981) might limit exposure in the hot tub to less than 15 minutes in 39 degree C water and less than 10 minutes in 40 degree C water. This recommended maximum time would be reduced if there are other risk factors present, e.g., the woman is not in good health, already has

an elevated body temperature from previously being in the hot tub or spa, fever, exercise, or another source of hyperthermia, or begins to feel uncomfortable.

The female consumer who may be pregnant and uses a hot tub or spa should be aware of these possible risks (Rogers and Davis, 1995). However, the current requirements for hot tub temperature control are not, in my judgement, sufficiently protective for use of these products by women who might be pregnant.

Although, in my judgement, the maximum allowable temperature that a hot tub can achieve need not be set below 38.9 degrees C, the female consumer should be able to monitor maximum water temperature to a level at or below a specified degree so that she can be assured that she will maintain her body temperature below 38.9 degrees C. Therefore, the product's temperature regulation system should ensure that the maximum water temperature achieved at a given setting is not exceeded due to significant variability in the mechanism that controls this level. The mechanism should be precise enough to ensure that a setting of 38.9 degrees C does not result in variability in the true temperature, i.e., that true temperature can be reliably and validly maintained. The current best recommendations are for women who might be pregnant to limit exposure to less than 15 minutes if the water temperature is at 38.9 degrees C. However, with the current standards for hot tub temperature regulation, a woman could limit her hot tub exposure according to these guidelines, but due to variability in the accuracy of the temperature regulation system, she could in reality be exposed to a true temperature of up to 41.9 degrees C.

And finally, the variability in water temperature within the tank should be taken into consideration, in that an individual who is positioned in the hot tub or spa near the inlet for newly heated water can be exposed to a much higher water temperature than that measured in other parts of the hot tub and reflected in the overall temperature gauge. Therefore, temperature gauges should be placed near the water inlet; and women who use hot tubs should be advised not to place themselves near the water inlet while using the hot tub so as to avoid exposure to water temperatures that exceed the overall recommended maximum.

VIII. References

- Andersen A-M M, Vastrup P, Wohlfahrt J, Andersen PK, Olsen J, Melbye M. (2002). Fever in pregnancy and risk of fetal death: a cohort study. *The Lancet*, 360: 1552-1556.
- Aoyama N, Yamashina S, Poelmann RE, Gittenberger-De Groot AC, Izumi T, Soma K, Ohwada T. (2002) Conduction system abnormalities in rat embryos induced by maternal hyperthermia. *The Anatomical Record*, 267: 213-219.
- Botto LD, Erickson JD, Mulinare J, Lynberg MC, Liu Y. (2002). Maternal fever, multivitamin use, and selected birth defects: evidence of interaction? *Epidemiology*, 13: 485-488.

Botto LD, Lynberg MC, Erickson JD. (2001). Congenital heart defects, maternal febrile illness, and multivitamin use: a population-based study. *Epidemiology*, 12: 485-490.

Cawdell-Smith J, Upfold J, Edwards M, Smith M. (1992). Neural tube and other developmental anomalies in the guinea pig following maternal hyperthermia during early neural tube development. *Teratogenesis, Carcinogenesis, and Mutagenesis*, 12: 1-9.

Centers for Disease Control and Prevention, <http://www.cdc.gov/ncbddd/>; accessed June 9, 2005.

Chambers CD, Johnson KA, Dick LM, Felix RJ, Jones KL. (1998). Maternal fever and birth outcome: a prospective study. *Teratology*, 58:251-257.

Chance PF and Smith DW. (1978). Hyperthermia and meningomyelocele and anencephaly. *The Lancet*, April 8, 769-770.

Christensen K. (2002). Vitamins, fever and birth defects – consistent interaction or persistent bias? *Epidemiology*, 13, 620-621.

Clarren SK, Smith DW, Harvey MAS, Ward RH. (1979). Hyperthermia – a prospective evaluation of a possible teratogenic agent in man. *Journal of Pediatrics*, 95: 81-83.

Edwards MJ. (1970). The experimental production of arthrogryposis multiplex congenital in guinea-pigs by maternal hyperthermia during gestation. *The Journal of Pathology*, 104: 221-229.

Edwards MJ, Gray CH, Beatson J. (1984). Retardation of brain growth of guinea pigs by hyperthermia: effect of varying intervals between successive exposures. *Teratology*; 29:305-12.

Edwards MJ, Shiota K, Smith MSR, Walsh DA. (1995). Hyperthermia and birth defects. *Reproductive Toxicology*, 9: 411-425.

Ferm VH and Ferm RR. (1979). Teratogenic interaction of hyperthermia and vitamin A. *Biology of Neonate*, 36: 168-172.

Finnell RH, Moon SP, Abbott LC, Golden JA, Chernoff GF. (1986). Strain differences in heat-induced neural tube defects in mice. *Teratology*, 33: 247-252.

Fisher NL and Smith DW. (1981). Occipital encephalocele and early gestational hyperthermia. *Pediatrics*, 68: 480-483.

Floyd RL, Decoufle P, Hungerford DW. (1999). Alcohol use prior to pregnancy recognition. *Am J Prev Med*; 17:101-7.

Germain M-A, Webster W, Edwards MJ. (1985). Hyperthermia as a teratogen: parameters determining hyperthermia-induced head defects in the rat. *Teratology*, 31: 265-272.

Graham JM, Edwards MJ, Edwards MJ. (1998). Teratogen update: gestational effects of maternal hyperthermia due to febrile illnesses and resultant patterns of defects in humans. *Teratology*, 58: 209-221.

Halperin LR and Wilroy, Jr. RS. (1978). Maternal hyperthermia and neural tube defects. *The Lancet*; 2:212-3.

Hartley WJ, Alexander G, Edwards MJ. (1974). Brain cavitation and micrencephaly in lambs exposed to prenatal hyperthermia. *Teratology*, 9: 299-304.

Harvey MAS, McRorie MM, Smith DW. (1981). Suggested limits to the use of the hot tub and sauna by pregnant women. *CMA Journal*, 125: 50-53.

Hellmann W. (1979). Effects of fever and hyperthermia on the embryonic development of rabbits. *The Veterinary Records*, April 28: 289-390.

Hendrickx AG, Stone GW, Henrickson RV, Mayayoshi E. (1979). Teratogenic effects of hyperthermia in the bonnet monkey (*macaca radiata*). *Teratology*, 19:177-182.

Kilham L and Ferm V. (1976). Exencephaly in fetal hamsters following exposure to hyperthermia. *Teratology*, 14: 323-326.

Kleinebrecht J, Michaelis H, Michaelis J, Killer S. (1979). Fever in pregnancy and congenital anomalies. *The Lancet*, June 30: 1403.

Kline J, Stein Z, Susser M, Warburton D. (1985). Fever during pregnancy and spontaneous abortion. *American Journal of Epidemiology*, 121:832-842.

Layde PM, Edmonds LD, Erickson JD. (1980). Maternal fever and neural tube defects. *Teratology*, 21:105-108.

Li D-K, Janevic T, Odouli R, Liu L. (2003a). Hot tub use during pregnancy and the risk of miscarriage. *American Journal of Epidemiology*, 158: 931-937.

Li H, Wang B, Zhao D, Han L. (2003b). Protective effects of vitamin B12, ginseng saponin, and folic acid against murine fetal deformities caused by hyperthermia. *Chinese Medical Journal*, 116: 1776-1778.

Lipson A. (1988). Hirschsprung disease in the offspring of mothers exposed to hyperthermia during pregnancy. *American Journal of Medical Genetics*, 29: 117-124.

- Lipson AH and Edwards MJ. (1993). Maternal sauna and hyperthermia during pregnancy and cardiovascular and other malformation in offspring. *European Journal of Epidemiology*, 9: 676-677.
- Little BB, Ghali FE, Snell LM, Knoll KA, Johnston W, Gilstrap LC. (1991). Is hyperthermia teratogenic in the human? *American Journal of Perinatology*, 8: 185-189.
- Lynberg MC, Khoury MJ, Lu X, Cocian T. (1994). Maternal flu, fever and the risk of neural tube defects: a population-based case-control study. *American Journal of Epidemiology*, 140: 244-255.
- Miller P, Smith DW, Shepard T. (1978). Maternal hyperthermia as a possible cause of anencephaly. *The Lancet*, March 11: 519-521.
- Milunsky A, Ulcickas M, Rothman KJ, Willett W, Jick S, Jick H. (1992). Maternal heat exposure and neural tube defects. *JAMA*, 268: 882-885.
- Rogers J and Davis BA. (1995). How risky are hot tubs and saunas for pregnant women? *MCN*, 20: 137-140.
- Sandford MK, Kissling GE, Joubert PE. (1992). Neural tube defect etiology: new evidence concerning maternal hyperthermia, health and diet. *Developmental Medicine and Child Neurology*, 34: 661-675.
- Sasaki J, Yamaguchi A, Nabeshima Y, Shigemitsu S, Mesaki N, Kubo T. (1995). Exercise at high temperature causes maternal hyperthermia and fetal anomalies in rats. *Teratology*, 51: 233-236.
- Saxen L, Holmberg PC, Nurminen M, Kuosma E. (1982). Sauna and congenital defects. *Teratology*, 25: 309-313.
- Shin J-H and Shiota K. (1999). Folic acid supplementation of pregnant mice suppresses heat-induced neural tube defects in the offspring. *Journal of Nutrition*, 129: 2070-2073.
- Shiota K, Shionoya Y, Masako I, Uenobe F, Kuwahara C, Fukui Y. (1988). Teratogenic interaction of ethanol and hyperthermia in mice. *Proceedings of the Society for Experimental Biology and Medicine*, 187: 142-148.
- Smith DW, Clarren SK, Harvey MAS. (1978). Hyperthermia as a possible teratogenic agent. *Journal of Pediatrics*, 92: 878-883.
- Tikkanen J and Heinonen OP. (1991). Maternal hyperthermia during pregnancy and cardiovascular malformations in the offspring. *European Journal of Epidemiology*, 7: 628-635.

Uhari M Mustonen A, Kouvalainen K. (1979). Sauna habits of Finnish women during pregnancy. *BMJ*, May 5: 1216.

Webster WS, Germain M-A, and Edwards MJ. (1985). The induction of microphthalmia, encephalocele, and other head defects following hyperthermia during the gastrulation process in the rat. *Teratology*, 31: 73-82.

Wilson JG and Fraser FC, Eds. *Handbook of Teratology*. New York: Plenum Press, 1977.